

## Les actes de colloques du musée du quai Branly Jacques Chirac

6 | 2016

Des êtres vivants et des artefacts

---

# Laisser vivre les machines

Dominique Lestel

---



### Édition électronique

URL : <http://journals.openedition.org/actesbranly/663>

ISSN : 2105-2735

### Éditeur

Musée du quai Branly Jacques Chirac

### Référence électronique

Dominique Lestel, « Laisser vivre les machines », *Les actes de colloques du musée du quai Branly Jacques Chirac* [En ligne], 6 | 2016, mis en ligne le 20 janvier 2016, consulté le 08 septembre 2020.

URL : <http://journals.openedition.org/actesbranly/663>

---

Ce document a été généré automatiquement le 8 septembre 2020.

© Tous droits réservés

---

# Laisser vivre les machines

Dominique Lestel

---

« Sans me vanter, je ne suis pas très intelligent  
non plus »

Haruki Murakami, 2003, *Kafka sur le rivage*, 10/18,  
p. 474.

- 1 Les Occidentaux appréhendent aujourd'hui la question du vivant à partir d'une double démarche : ils cherchent d'abord à établir les frontières entre ce qui est vivant et ce qui ne l'est pas et ils tentent ensuite de trouver une définition objective du vivant en termes de conditions nécessaires et suffisantes<sup>1</sup>. Une telle façon de procéder nous paraît naturelle parce que nous sommes devenus de vrais cartésiens, mais elle reste très problématique. Elle suppose tout d'abord qu'il existe une classe d'éléments (les êtres vivants) qui nous serait spontanément accessibles de façon globale et que nous pourrions l'étudier (l'observer et l'expérimenter) pour en déterminer objectivement les propriétés. Elle postule ensuite qu'il suffirait d'en établir la synthèse et d'en tirer quelques règles simples à partir desquelles nous pourrions déterminer si nous avons affaire à du vivant ou non. La démarche cartésienne de l'identification suivie par une définition exhaustive et conclue par la reconnaissance des éléments qui appartiennent à l'ensemble en question se heurte pourtant à de très nombreuses difficultés – théoriques, comme celle de savoir comment déterminer la pertinence des critères retenus ; pratiques, comme celle de savoir si des entités ambivalentes comme les virus ou les cristaux sont vivants ou non et logiques comme celle de savoir si on peut légitimement restreindre le vivant à l'ensemble de ce que nous considérons comme tel dans notre culture.
- 2 La robotique sociale (c'est-à-dire la conception et la construction de robots capables d'avoir des interactions sociales avec des humains, des animaux ou d'autres robots) constitue un exemple intéressant dans un débat sur les frontières entre ce qui est vivant et ce qui ne l'est pas et sur l'imbrication des êtres vivants et des artefacts dans les nouvelles technologies de l'information et de la cognition qui constitue le sujet de ce colloque du Musée du Quai Branly (Coupaye 2016 ; Pitrou 2016). Le problème réside dans le conflit cognitif dans lequel nous sommes pris entre le désir de considérer ces artefacts perturbants comme des agents vivants et une conception hégémonique du

vivant dans les cultures occidentales qui fait des biologistes les experts patentés du sujet. Je vais proposer ici une conception relationnelle et constructiviste du vivant qui sort du schéma intellectuel habituel occidental. Dans cette perspective, est vivant ce qu'un groupe humain est prêt à considérer comme tel à travers des dispositifs technico-culturels, des postures cognitives, et une façon sophistiquée de manipuler le désir de vivant dans des mises en scène performatives plus ou moins élaborées. Il ne s'agit pas de prétendre que n'importe quoi peut être considéré comme vivant si on le veut, mais de comprendre qu'un phénomène comme celui du vivant recèle des ressources encore insoupçonnées, est infiniment plus complexe que ce qu'on est encore prêt à entendre et entre difficilement dans les catégories dépourvues d'ambiguïté d'une logique cartésienne contemporaine des « jardins à la française » taillés au cordeau. Un dernier point important avant de commencer : le texte qui suit est un essai philosophique qui ne cherche pas à prouver ou à démontrer quoi que ce soit. Son objectif est de déployer les conditions de possibilité d'un phénomène (ici le vivant), d'en explorer les conséquences les plus intéressantes et de conceptualiser au mieux ce qui peut l'être.

## James et Turing, pour commencer

- 3 Des cultures différentes donnent des réponses sensiblement divergentes à cette question de savoir si des artefacts peuvent être vivants. En Occident, l'élite intellectuelle répond habituellement d'une façon plutôt négative en estimant qu'une chose ou une machine n'est pas vivante. On peut sans doute les imaginer ainsi, mais elles ne le sont pas *vraiment*. L'arrivée de robots et de machines computationnelles incite à reposer la question et à donner des réponses apparemment plus nuancées. Je dis « apparemment » parce que ces réponses sont en fait très problématiques, même quand elles s'évaluent elles-mêmes comme raisonnables. Contrairement à ce qu'il imagine, la réponse de l'ingénieur positiviste qui estime que ses robots ne sont *que* des machines, au même titre qu'un grille-pain, n'est pas plus satisfaisante que celle du philosophe analytique qui demande que les Droits de l'Homme puissent s'appliquer aux robots. Avec certains sujets, les réponses raisonnables se révèlent finalement tout à fait délirantes<sup>2</sup>. Je vais avancer moi-même un point de vue un peu particulier sur le statut de vivant des artefacts et essayer de le défendre. Je vais commencer par évoquer deux textes philosophiques qui me semblent intéressants pour comprendre ce qui est en jeu avec les robots sociaux. Le premier est une note du philosophe américain William James sur les raisons pour lesquelles un homme ne pourra jamais tomber amoureux d'une femme artificielle. Le second est l'article princeps du pionnier britannique de l'informatique, Alan Turing, dans lequel il expose son célèbre jeu de l'imitation – un article tellement riche que son interprétation est loin d'avoir été épuisée.
- 4 Le texte de William James est publié en 1908 (James 1909). Dans une petite note concise, William James établit la différence fondamentale qui sépare une femme artificielle qui serait *exactement* comme une vraie femme et cette dernière : la première ne peut pas se soucier réellement *de moi* – et encore moins se soucier de moi positivement – ce qui m'empêche d'en tomber amoureux. Au mieux, cette femme artificielle ne peut exhiber que des symptômes qui ne renvoient à aucune conscience réelle. Pour James, l'artefact se distingue de la vraie femme par son incapacité à se soucier de moi – c'est-à-dire à faire preuve de sympathie et de reconnaissance, d'amour et d'admiration. Le souci de moi n'est pas non plus exactement le souci de soi, comme le lecteur va très bientôt s'en

rendre compte. Bien sûr, le souci de moi n'est pas exactement ce dont parle James. Pour lui, la différence est plutôt que la femme artificielle ne me donne pas l'opportunité de croire qu'elle se soucie de moi. Le point intéressant de la remarque de James est qu'il pense une situation comme celle de tomber amoureux à partir d'une posture du point de vue de la 2<sup>e</sup> personne. L'état amoureux s'établit à partir de ma représentation des représentations que l'autre a de moi. La difficulté qu'il rencontre est que le problème de la croyance, qu'il mobilise à cette occasion, renvoie à quelque chose d'extrêmement difficile à objectiver. De ce point de vue, l'intérêt du « test de Turing » est de mettre en place un dispositif qui me permette de substituer une situation interactionnelle constamment améliorable à un état psychologique, celui de la croyance, inconsistant et terriblement délicat à saisir de façon satisfaisante.

- 5 Le texte dans lequel Turing expose son fameux test (destiné à établir qu'une machine est intelligente) est publié en 1950 (Turing 1950). Quoique le test de Turing ait été abondamment commenté au cours des années, son importance a été très largement sous-estimée, même par ceux qui s'y sont référés avec passion. Le « test de Turing » propose en effet une mise en scène conceptuelle très féconde en adoptant une approche cooptative de la catégorie, sensiblement différente des approches logiques habituelles. Au départ, Turing construit une épreuve comparative autour du genre. Un homme ou une femme peut-il/peut-elle distinguer un homme d'une femme, uniquement à travers leurs réponses à ses questions ? Très vite une autre question se substitue à la première : un humain est-il capable de distinguer un humain en lui posant uniquement des questions ? L'interrogation prend d'ailleurs une forme sensiblement différente : un ordinateur peut-il tromper un humain sur sa qualité à travers la seule communication symbolique ? Le texte de Turing insiste sur la nécessité de passer par un support purement symbolique : « nous jugerions sans intérêt de rendre une 'machine pensante' plus humaine en l'habillant d'une... chair artificielle ». Turing précise que le test doit se faire dans des conditions d'abstraction radicale, « à travers les conditions qui empêchent l'interrogateur de voir ou de toucher les autres participants, ou d'entendre leur voix ». Le test de Turing porte sur une situation de handicap autant que sur une situation de capacité. Évoquant l'éducation d'un programme, il dira que ce dernier n'a ni bras ni jambes ni perceptions comme un vrai enfant, et il se réfère à Helen Keller<sup>3</sup> pour évacuer un examen sérieux de la question : *pense-t-on aussi avec les jambes ?* C'est le paradoxe central du test de Turing : évaluer la capacité communicationnelle d'un agent en le plaçant dans une situation de pauvreté communicationnelle extrême. Malgré les handicaps massifs auxquels s'est astreint Turing, la conception de son test s'appuie sur plusieurs tournants majeurs dans la façon de voir l'intelligence<sup>4</sup> : passage de la définition de l'intelligence à celle de sa reconnaissance ; approche cooptative de l'intelligence ; la ruse et la tromperie au centre de ce qu'est l'intelligence et une conception de l'intelligence du point de vue de la 2<sup>e</sup> personne. Ce dernier point doit être d'emblée précisé. L'une des ambiguïtés de l'article de Turing se trouve en effet dans le statut attribué à la machine avant que celle-ci soit identifiée comme intelligente/non intelligente. Qu'est-ce que ce « elle » ? La posture de l'intelligence du point de vue de la 2<sup>e</sup> personne résout le problème. La machine ne dit jamais « je ». Le processus est toujours parfaitement situé : c'est l'humain qui rend toujours compte du processus. On n'est pas *dans* un dialogue mais dans un monologue à deux... jusqu'à ce que la machine se révèle intelligente à *travers* le dialogue. Une question que soulève le test de Turing est celle de sa généralisation possible en l'appliquant au vivant.

## Une autre approche du vivant

- 6 Ces deux textes adoptent une démarche très différente de celle habituellement suivie à propos des questions dont ils traitent – la capacité de tomber amoureux d'un artefact pour James, celle de se faire passer pour un humain quand on est une machine pour Turing. Bien que ni James ni Turing n'abordent jamais explicitement la question de ce qu'est la vie, ils ouvrent la voie à une démarche féconde et originale pour la penser de façon inédite à travers une approche concrète, relationnelle, perspectiviste et constructiviste de ce qu'est un être vivant. À la différence de l'approche réaliste, objectiviste, causaliste et essentialiste qui est majoritairement adoptée par les chercheurs et les philosophes, une place est laissée à l'étude de certains artefacts perturbants.

## Le vivant du point de vue de la première personne

- 7 Les deux textes évoqués conduisent à l'idée selon laquelle la question de ce qui est vivant *pour moi* est peut-être une question plus pertinente que celle de savoir ce qui est vivant de façon plus générale. Dans cette perspective, ce qui est vivant pour moi est ce qui peut partager *mon* existence. L'idée qu'il existerait une catégorie unique qui contiendrait tout ce qui est vivant et qu'une telle catégorie pourrait être objectivée, c'est-à-dire qu'elle pourrait être définie sans aucune référence à un être vivant concret qui le caractériserait apparaît très problématique compte tenu de la multiplicité des occurrences à travers lesquelles on est prêt à reconnaître le vivant comme vivant. On peut légitimement se demander si ce qui pose un problème n'est pas l'ambition universaliste qui sous-tend la recherche d'une définition générale du vivant qui s'appliquerait mal à la multiplicité de ses occurrences possibles. Turing et James attirent justement l'attention sur le fait que quand quelqu'un identifie un agent comme ayant une propriété un peu complexe, comme l'amour, l'intelligence ou la vie, il fait lui-même pleinement partie du processus. La question à la troisième personne : « Qu'est-ce qu'un être vivant ? » gagne donc à se formuler à la première personne : « Qu'est-ce qu'un être vivant *pour moi* ? », ce que ne font jamais les biologistes ou ceux qui aspirent à l'être comme les ingénieurs de la Vie Artificielle. L'avantage d'une telle posture est que la nature de l'agent à évaluer (qu'il soit naturel ou artificiel ne change pas grand-chose à l'histoire) n'a plus aucune importance. L'ensemble des êtres vivants est alors constitué de l'ensemble de ceux qui sont considérés comme des êtres vivants pour au moins un autre être vivant. Pour adapter les textes de Turing et James, est un être vivant ce que je reconnais comme tel (Turing) et n'est pas un être vivant ce que je ne reconnais pas comme tel (James). Les deux points de vue sont complémentaires et superficiellement redondants. Pour l'exprimer d'une autre façon, dans le dispositif intellectuel de Turing, peut-être considéré comme vivant un agent que les experts patentés du vivant (les biologistes) ne sont pas prêts à admettre dans le club des êtres vivants mais que je suis moi-même prêt à admettre comme être vivant dans ma vie. Et dans la perspective de James, ne sera pas considéré comme vivant un agent que les biologistes accréditent totalement mais qui ne me convainc pas pour autant et que je ne suis donc pas prêt à accepter comme vivant dans ma propre vie. En termes plus

brutaux, il existe d'autres façons de concevoir le vivant que celle que proposent les biologistes et nous avons tout intérêt à les approfondir.

## Une conception du vivant à la 2<sup>e</sup> personne

- 8 Mais si on les lit plus attentivement encore, les textes de James et Turing proposent moins une conception du vivant à la première personne (est vivant ce que je peux considérer être vivant) qu'une conception du vivant à la 2<sup>e</sup> personne, (je te considère comme vivant si je peux *t'éprouver* comme vivant dans la confrontation que j'ai avec toi et si je peux t'attribuer une volonté/un désir ou tout simplement un point de vue qui se confronte avec le mien). Dans le point de vue à la deuxième personne, le sujet se constitue à travers ses frictions avec des interlocuteurs qui le considèrent lui-même comme interlocuteur. Une telle notion est discutée en philosophie par quelques rares penseurs tels que Martin Buber (1937) (l'individu émerge dans la relation Je-Tu), Francis Jacques (l'intersubjectivité prime sur la subjectivité dans le phénomène du dialogique et la présence de l'absence est essentielle dans ce processus) et Georges Herbert Mead (le soi ne se connaît lui-même qu'en se mirant dans un *alter ego* à la deuxième personne). Tous ces auteurs ont cependant restreint leur approche au dialogue verbal, d'une part, et à la communication entre humains (et éventuellement entre humains et Dieu), d'autre part<sup>5</sup>. Envisager que des êtres vivants non humains puissent adopter eux-mêmes un point de vue à la deuxième personne en élargit considérablement la portée. Passer du dialogue verbal à la confrontation des points de vue aussi.

## Paradigme constructiviste

- 9 Une telle approche relationnelle du point de vue de la 2<sup>e</sup> personne acquiert une certaine pertinence pour penser la robotique sociale si on la place dans le paradigme du *constructivisme* – qui diffère sensiblement du relativisme auquel il est parfois abusivement assimilé – qui lui donne son sens. Les êtres vivants abstraient des régularités et des règles qui leur permettent d'éviter les expériences désagréables et de générer des situations agréables. Pour le constructivisme, nous ne pouvons pas construire toutes les réalités que nous voulons, car il faut que nos constructions soient *viabiles* même si elles ne sont pas *vraies*. Le savoir n'est pas une image de la réalité, mais un répertoire des actions et des pensées qui se sont révélées être efficaces dans l'expérience passée. Comme un alpiniste chevronné, vous devez d'abord vous représenter ce que vous ne pouvez pas faire. Vous voyez les avalanches possibles, les crevasses, etc. Vous faites des choix mais à l'intérieur de l'espace laissé entre les contraintes de la montagne. Leibniz et Maupertuis ont donné une bonne formulation de la viabilité avec le principe de moindre action ou de moindre résistance. Une activité comme le jeu d'échec génère bien des certitudes, mais celles-ci n'ont de sens que dans le cadre très restrictif du jeu. Pour le constructivisme radical, la connaissance ne reflète pas une réalité ontologique "objective", mais la mise en ordre et l'organisation d'un monde qui se constitue à travers notre expérience. Comme le dit Piaget (1937 : 311), l'intelligence organise le monde en s'organisant elle-même. Une telle idée s'éclaire très bien à travers la distinction élaborée par von Glasersfeld entre *convenance* (to fit) et *correspondance* (to match). La relation entre connaissance et réalité est vue dans l'épistémologie traditionnelle et la psychologie cognitive la plus répandue comme une

correspondance plus ou moins figurative (iconique), alors que le constructivisme la conçoit comme une adaptation au sens fonctionnel. Les cambrioleurs professionnels savent « qu'il existe beaucoup de clefs découpées tout à fait différemment des nôtres, mais qui n'en ouvrent pas moins nos portes ». Et von Glasersfeld ajoute : « Nous nous trouvons face à notre environnement comme le cambrioleur face à une serrure qu'il doit ouvrir pour atteindre le butin qu'il espère emporter ». C'est pourquoi l'anthropologue Gregory Bateson attribue une place privilégiée à la cybernétique parce qu'elle n'explique pas au moyen de relations causales mais en spécifiant des contraintes. Il conçoit la cybernétique comme art de créer des équilibres dans un monde de possibilités et de contraintes. Pour Von Glasersfeld (1995), la cybernétique fournit un style de vie et offre un nouvel impératif éthique : générer de nouvelles possibilités plutôt que de frapper nos têtes contre des contraintes que nous ne pouvons changer. Dans le schéma causaliste, la liberté n'existe pas, ni l'invention. Dans l'approche constructiviste, en dehors des contraintes imposées, tout est possible et tout ce qui n'est pas compatible avec les contraintes données est éliminé. Ce qui permet à von Glasersfeld d'écrire : « le principe fondamental de l'épistémologie constructiviste radicale coïncide avec celui de la théorie de l'évolution » (Von Glaserfeld 1988).

## Le vivant comme contrat existentiel

- 10 J'ai proposé plus haut une lecture constructiviste de James et de Turing. Le mathématicien anglais met en place un dispositif élaboré pour faire émerger la nature de l'artefact là où le philosophe pragmatiste n'envisage qu'une confrontation interindividuelle un peu sommaire. Pour Turing, la reconnaissance de l'intelligent est concomitante d'une mise en scène un peu élaborée qu'il assume totalement. La reconnaissance de l'intelligent passe par une mise en demeure savamment orchestrée dans un espace concurrentiel dans lequel les éléments pertinents qu'il est possible de prendre en compte ne sont pas limités *a priori*. L'intelligence de la machine n'est pas une propriété qu'un philosophe ou un scientifique un peu malin pourrait découvrir, mais un contrat existentiel qui se construit à partir des technologies disponibles et d'un désir de rendre vivant l'artefact.
- 11 Il ne suffit pas de vouloir avoir des machines vivantes pour le faire mais le désir de vouloir y parvenir est une condition nécessaire pour y parvenir. D'où l'intérêt de penser la robotique sociale contemporaine au croisement d'une conceptualisation de ce qui passe pour vivant et de ce qui pourrait être caractérisé comme des dispositifs libidinaux qui manipulent le désir de vivant et qui maintiennent captive la volonté de recevoir l'artefact dans le club des êtres vivants. Une telle attitude est très différente de celle du biologiste classique qui s'accrochera à des pré-requis non négociables, par exemple à la présence ou non d'ADN. L'agencement des cultures libidinales, des virtuosités technologiques et des pratiques du vivant permet ainsi de mieux comprendre ce qu'est un robot social qui à la fois simule *un* être vivant et constitue un piège pour capter notre désir de « faire vie » avec lui – simuler devant être pris ici au double sens qu'il a en français de « faire comme si » et de « faire prendre pour ». La vraisemblance<sup>6</sup> du robot découle ainsi non seulement de sa conception mais aussi de sa capacité à nous intoxiquer de notre souci pour lui et de son souci pour nous en disposant l'un et l'autre dans un dispositif de dépendance mutuelle. Un artefact devient vivant à l'intérieur d'un dispositif complexe (technique, social, sémiotique, psychologique) qui *me* le rend vivant, à condition que j'en éprouve le désir – un désir

qui est d'ailleurs loin d'être un donné élémentaire mais qui est plutôt déjà lui-même une construction élaborée. Une approche du vivant du point de vue de la 2<sup>e</sup> personne lie donc ma capacité à concevoir des agents dotés d'une certaine autonomie par rapport à mon désir de partager ma vie avec eux.

## Émergence des machines libidinales

- 12 Dans une perspective constructiviste et relationnelle du point de vue de la 2<sup>e</sup> personne, les robots sociaux doivent être considérées comme des machines libidinales qui suscitent et manipulent *mon* désir. La robotique sociale nous offre ainsi l'opportunité de repenser des notions fondamentales, comme celles de vie ou d'intelligence, en dehors des espaces convenus (ceux de la biologie ou de la psychologie). La robotique est de ce point de vue une non-biologie et une post-psychologie. C'est une non-biologie en ce sens qu'elle remet la biologie à sa place d'étude des constituants du vivant carboné, et non du vivant en général. C'est une post-psychologie, parce qu'elle n'a aucune raison de s'appuyer pas sur des catégories culturelles surdéterminées, comme celle de raisonnement ou celle de soi<sup>7</sup>, pour rendre compte de l'intelligence des machines. Ce qui est fécond n'est pas d'appliquer des catégories rebattues et de plus en plus problématiques aux artefacts quasi-autonomes contemporains mais de saisir l'opportunité des difficultés rencontrées pour sortir de ces catégories culturellement datées. En d'autres termes, la question n'est pas de savoir si les robots sont vivants mais quels sont les dispositifs que je peux mobiliser à cette fin et les pratiques dans lesquels je suis prêt à m'engager pour leur donner un tel statut. La robotique sociale, depuis ses débuts, s'oriente vers une captation un peu sauvage du désir de leur donner un tel statut. Dans son étude sur les jeux de hasard des casinos de Las Vegas, la sociologue Natasha Dow Schull a démonté les dispositifs techniques qui manipulent les désirs du joueur en lui faisant croire en une autonomie illusoire et en accentuant son sentiment qu'il maîtrise un processus qui lui échappe en fait totalement (Dow Schull 2005).
- 13 Dès leurs origines, les robots sociaux génèrent et manipulent le désir ; ils le suscitent et sont conçus pour y répondre<sup>8</sup>. Les réduire à de pures fonctionnalités est très appauvrissant et permet difficilement de comprendre ce qui est vraiment en jeu avec elles. C'est pourquoi il est particulièrement pertinent de s'intéresser aux premières fictions qui portent sur des robots sociaux (Hoffmann, Villiers de L'Isle-Adam, etc.) et qui évoquent explicitement des machines qui suscitent le désir. Doit-on rappeler que le problème que rencontre Frankenstein est précisément que sa créature veut connaître l'amour<sup>9</sup>? Nul déterminisme n'a imposé cette orientation. D'autres cultures ont fait d'autres choix – par exemple en plaçant l'artefact intelligent dans l'espace de la médiation aux dieux, aux esprits et aux démons. La sensibilité artistique et littéraire se distingue là singulièrement des blocages de la pensée savante – et la sociologue Sherry Turkle qui s'émeut qu'on puisse faire dire à un robot « je t'aime » passe à côté de quelques-uns des enjeux majeurs de la robotique parce qu'elle reste dans un espace intellectuel trop conventionnel pour ça – alors même qu'elle fut l'une des premières à mettre l'accent sur la dimension identitaire des nouvelles technologies de l'information (Turkle 2013).
- 14 Le statut de ces robots n'est donc pas aussi tranché qu'on pourrait l'imaginer. Il s'agit moins de déterminer s'ils sont vivants ou non que de reconnaître leur ambivalence fondamentale dans l'espace du vivant. La question est infiniment plus large et riche que



celle de savoir si des robots peuvent par exemple être déclarés « responsables » ou si j'ai des devoirs éthiques vis-à-vis de certains d'entre eux. Il est tout aussi illusoire de vouloir penser ces artefacts dans les catégories convenues à travers lesquelles on appréhende habituellement les technologies et l'histoire des technologies. Il me semble important de penser l'appartenance des robots à une catégorie d'artefacts que négligent généralement les histoires des techniques et la majorité des philosophes des techniques mais qu'évoquent constamment les anthropologues, les ethnopsychiatres et les historiens de l'art – la catégorie des *artefacts générateurs de sens*. Des artefacts qui ne servent pas à faire quelque chose mais qui contribuent à donner du sens et une présence à un espace ou à une situation. Des artefacts qui génèrent des existences<sup>10</sup> ou des présences plutôt que des procédures, même s'ils peuvent aussi générer des procédures. Ce sont à la fois des artefacts qui ont un rapport particulier au vivant et des artefacts qui donnent une vision particulière du vivant. En faisant de la robotique sociale une question d'ingénierie de substitution du vivant, nous nous égarons; les robots sont plutôt des artefacts qui nous obligent à repenser ce que peut être le vivant et c'est ce qui les rend si intéressants pour nous. Le phénomène des artefacts vivants n'émerge pas avec les robots ou les ordinateurs, comme le croient encore la majorité des théoriciens contemporains qui s'intéressent à la question. Il s'inscrit au contraire dans une histoire d'une très grande richesse et d'une immense complexité – dans lequel on compte les marionnettes, certes, mais aussi les fétiches, les gris-gris et les amulettes. L'une des caractéristiques centrales de ces artefacts est de ne pouvoir fonctionner qu'à l'intérieur de dispositifs de performances, dont quelques-unes ont été bien étudiées par des anthropologues.

## Les performances de la robotique sociale

- 15 La posture constructiviste et relationnelle nous évite de devoir choisir entre une vision scientifique essentialiste et une vision relativiste de convenance : aucune des deux perspectives n'est satisfaisante. Une des difficultés que nous éprouvons à penser la robotique sociale trouve ses racines dans une organisation universitaire trop bureaucratique des savoirs. Nos catégories bien tranchées entre ingénierie, science, philosophie, religion, art et littérature appréhendent difficilement la complexité de ce qui est en jeu avec ces robots qui se personnalisent. La robotique sociale doit plutôt être pensée au croisement de champs disciplinaires multiples qui sont habituellement pensés comme disjoints.
- 16 Dans l'approche constructiviste et relationnelle à laquelle je me réfère ici, le statut des robots n'est pas donné d'emblée. Il est progressivement élaboré en fonction des choix existentiels que nous allons adopter et des dispositifs à travers lesquels nous allons leur laisser une place. Rendre vivants quelques-uns de nos artefacts ne signifie pas seulement les transformer techniquement pour les faire entrer dans une norme préétablie mais aussi nous transformer nous-mêmes pour rendre nos existences mutuellement compatibles – notre existence et celles de certains artefacts. L'enjeu des robots sociaux, en d'autres termes, est de créer des communautés hybrides de partage de sens, d'intérêts et d'affects<sup>11</sup> avec des agents dont le statut même de vivant fait partie de ce qui doit être négocié au sein de la communauté, sachant que ce n'est jamais la seule chose qui y est négociée. Avec les robots, c'est la question de l'existence plus que celle de la vie *sensu stricto* qui est en jeu. Une façon de se rendre compatible est de s'

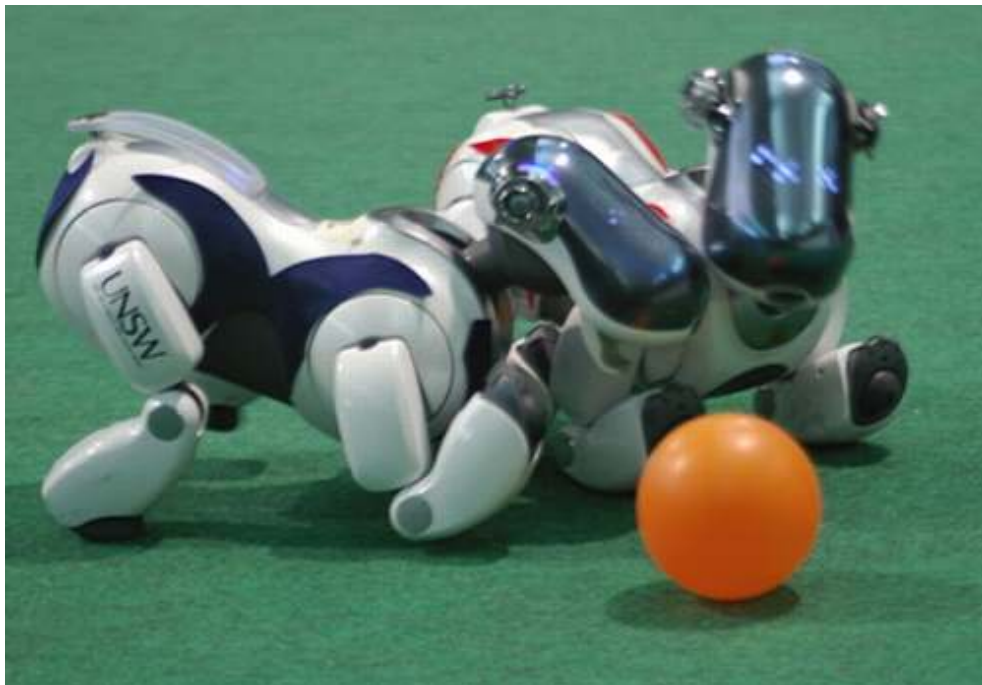
*enmachiner*. Une façon plus féconde et plus riche de le faire est de mobiliser des fictions engageantes qui nous conduisent à reconceptualiser la porosité de l'espace du vivant et à comprendre le défi posé par les robots comme l'opportunité de repenser la place de l'animisme dans des cultures techno-saturées<sup>12</sup>. Il y en a sans doute d'autres. Dans une approche constructiviste du monde, être vivant n'est pas une *propriété* mais un dispositif relationnel et performatif qui met en jeu des compétences ou des capacités, sans s'y réduire pour autant ni en avoir la moindre exclusivité. C'est précisément ce à quoi renvoie le titre de cette présentation : *laisser vivre les machines*. Être vivant engage une *posture passive* autant qu'elle mobilise une *capacité active*. Les Occidentaux ont beaucoup de difficultés à comprendre cette intuition parce qu'ils évoluent dans une culture qui dénigre la passivité au profit de l'action. Les marionnettes et les robots sociaux ne deviennent des êtres vivants que si nous leur en donnons l'opportunité, si nous leur offrons une écologie dans laquelle elles peuvent vivre, et si nous arrivons à établir entre elles et nous une forme de non intervention positive<sup>13</sup>. Une telle attitude est forcément problématique dans une culture occidentale où la technologie est vécue avant tout comme espace d'intervention active. Mais la capacité de rendre l'artefact vivant provient d'une capacité à lui accorder l'hospitalité dans *notre* monde vivant, à nous transformer pour accorder un statut de vivant à « un autre », et à rendre nos vies mutuellement compatibles. L'idée importante ici, c'est que le statut de vivant que peut recevoir un artefact ne dépend pas uniquement de ce qu'il est ou de ce qu'il fait, mais de la façon dont nous nous voyons nous-mêmes et de la façon dont nous sommes prêts à nous engager avec lui pour parvenir à un tel résultat. Les efforts, en d'autres termes, sont partagés. C'est précisément là que la question du point de vue de la 2<sup>e</sup> personne devient cruciale car l'artefact acquiert un statut de vivant par la capacité qu'il a à nous engager avec lui comme s'il était vivant et à le faire de telle sorte qu'il devient précisément vivant. Ce qui rend un Paro (fig. 1)<sup>14</sup> vivant ce n'est pas qu'il dit qu'il nous aime mais que nous l'aimons nous-mêmes à travers les marques d'affection qu'il nous prodigue et celles qu'on lui donne en retour. Il devient vivant parce que nous le sentons vivant, et nous le sentons vivant parce qu'il nous engage vis-à-vis de lui dans une telle relation. Nous construisons avec lui un espace dans lequel le fait qu'il soit en vie a un sens.

Figure 1. Le robot Paro.



Photographie : Theron Trowbridge

Figure 2. Le robot AIBO.



Photographie : Brad Hall

- 17 Le point important à saisir ici, c'est que dans une perspective relationnelle et constructiviste il n'y a pas de différence entre l'affection et la simulation de l'affection, et que cette simulation de l'affection peut atteindre une force de conviction réellement très intense, si l'on mobilise les dispositifs performatifs et relationnels *ad hoc* pour y parvenir. Un esprit positiviste objectera que l'on est dans l'illusion, sauf que quand un

enfant devient extrêmement malheureux parce qu'on veut remplacer son Aibo<sup>15</sup> (fig. 2) habituel par un autre Aibo, exactement pareil au premier, il ressent des émotions négatives qui n'ont rien de simulées et qui l'affectent comme n'importe quelles autres émotions négatives. La clef de ce qui se passe réside dans un phénomène très simple et pourtant sous-estimé : la non transitivité des simulations d'affects – je peux ressentir des affects très intenses à partir d'autres affects purement simulés. Même si le robot n'a que des affects simulés, je peux l'introduire dans un piège technico-psychique qui génère en moi des affects que je ne suis pas prêt à accepter comme simulés et qui me conduisent à reconnaître au robot un statut qui n'est peut-être pas le même que celui d'un chien mais qui est assurément très différent de celui que j'attribue à un grillean. En d'autres termes, je me mets en position de reconnaître le robot comme vivant à travers la reconnaissance qu'il a de moi. Je lui attribue sinon *la* vie, du moins *une* vie, à travers une posture du point de vue de la 2<sup>e</sup> personne qui se concrétise à travers un dispositif technique contraignant que j'ai construit à cette fin. On peut alors se demander pourquoi j'ai besoin d'avoir des artefacts auxquels j'attribue ainsi un statut de vivant, mais c'est assurément une autre question. J'obtiens ainsi une forme de vivant qui échappe à l'espace de prédation du biologiste. Ce n'est pas parce qu'ils sont vivants que nous pouvons nous comporter avec eux comme avec des êtres vivants, mais bien parce que nous nous comportons avec eux comme s'ils étaient des êtres vivants qu'ils le deviennent en fin de compte.

- 18 L'espace des catégories mobilisables pour parler du vivant doit être repensé pour fournir un espace opérationnel pour ces fictions engageantes qui nous constituent en même temps qu'elles établissent la nature de ceux avec qui on partage notre existence et qui déstabilisent la notion même d'existence en la faisant déborder en dehors de ses espaces sémantiques familiers. Quelques-unes de ces fictions apparaissent au premier abord très étranges mais elles atteignent parfois une pertinence jusque-là insoupçonnée. Certains agents, comme les robots sociaux, vivent dans des vies qui ne sont pas biologiques, et qui se développent par contamination plutôt que par reproduction. C'est un phénomène qui mérite d'attirer notre attention et d'inquiéter nos visions du monde un peu trop cartésiennes.

---

## BIBLIOGRAPHIE

Anders, Günther

1957/2002 *L'obsolescence de l'homme. Sur l'âme à l'époque de la deuxième révolution industrielle*. Paris, Éditions de l'Encyclopédie des Nuisances.

Bardini, Thierry

2000 *Bootstrapping : Douglas Engelbart, coevolution, and the origins of personal computing*. Stanford, Stanford University Press.

Buber, Martin

1937 *I and Thou* (traduit par Ronald Gregor Smith), Edimbourg, T. and T. Clark, 2<sup>e</sup> édition, New

York : Scribners, 1958. 1st Scribner Classics éd. New York, NY : Scribner, 2000, 1986. Publication originale 1923 *Ich und du*.

Coupaye, Ludovic

2016 « Yams as vernacular methodology? Approaching vital process through technical processes », in Perig Pitrou, Ludovic Coupaye et Fabien Provost (éds.), *Actes du colloque « Des êtres vivants et des artefacts : l'imbrication des processus vitaux et des processus techniques », 9 et 10 avril 2014, musée du quai Branly, Paris*. [<https://actesbranly.revues.org/673>].

Davis, Erik

1998 *Technognosis : myth, magic + mysticism in the age of information*. New York, Three Rivers Press.

Dow Schull, Natasha

2005 « Digital gambling : the coincidence of desire and design », *Annals of the American Academy of Political and Social Science* 597 : 65-81.

Grimaud, Emmanuel et Zaven Paré

2011 *Le jour où les robots mangeront des pommes*. Paris, Pétra.

James, William

1909 *The meaning of truth : a sequel to "pragmatism"*. Chapitre VIII, originellement publié dans : *Philosophical Review*, Janvier 1908 xvii (1).

Lecerle, Jean-Jacques

1988 *Frankenstein : mythe et philosophie*. Paris, PUF.

Lestel, Dominique

1996 *L'animalité : essai sur le statut de l'humain*. Paris, Hatier. Réédité chez Les Cahiers de l'Herne en 2004.

Lestel, Dominique

1996 « Metaphors of complexity : language and cognitive resources of artificial life », *Social Science Information* 35 (5) : 511-540.

Lestel, Dominique, Louis Bec et Jean-Louis Lemoigne

1993 « Visible characteristics of living systems : esthetics and artificial life », in J. L. Deneubourg et al. (éds.), *Self-organization and life : from simple rules to global complexity*. Bruxelles, Proceedings of the European Conference on Artificial Life : 595-603.

Murakami, Haruki

2003 *Kafka sur le rivage*. Paris, 10/18.

Piaget, Jean

1937 *La construction du réel chez l'enfant*. Neuchâtel, Delachaux et Niestlé.

Pitrou, Perig

2016 « Êtres vivants/artefacts, processus vitaux/processus techniques : remarques à propos d'un cadran analytique », in Perig Pitrou, Ludovic Coupaye et Fabien Provost (éds.), *Actes du colloque « Des êtres vivants et des artefacts : l'imbrication des processus vitaux et des processus techniques », 9 et 10 avril 2014, musée du quai Branly, Paris*. [<https://actesbranly.revues.org/653>].

Turing, Alan

1950 « Computing machinery and intelligence », *Mind* 59 (236) : 433-460.

Turkle, Sherry

2013 *Alone together*. New York, Random House.

Turner, Fred

2006 *From counterculture to cyberculture : Stewart Brand, the whole earth network and the rise of digital utopianism*. Chicago, University of Chicago Press.

Von Glasersfeld, E.

1995 « A cybernetician before cybernetics », présenté lors de la rencontre annuelle de l'American Society for Cybernetics. Publié dans 1997, *Systems Research and Behavioral Science* 14 (2) : 137-139.

Von Glasersfeld, E.

1988 « Introduction à un constructivisme radical », in P. Watzlawick (éd.), *L'invention de la réalité*. Paris, Seuil.

## NOTES

1. Je tiens à remercier le CNRS qui m'a accordé une délégation à l'Université de Tokyo en 2013-2014 pour travailler sur ces questions, Sandra Laugier, sans qui rien de ceci n'aurait été possible, Philippe Codognot pour son hospitalité au sein du *Japanese French Laboratory for Informatics* et Perig Pitrou pour m'avoir invité à Paris pour intervenir au musée du quai Branly et avoir fait une lecture critique du texte qui a permis de sensiblement en améliorer la qualité.
2. A ce sujet, voir Günther Anders (1957/2002 : 34-35) : « Notre exposé 'outrancier' n'est donc qu'un reflet partiel de l' 'exagération' qui a réellement lieu aujourd'hui : *il n'est que l'exposé outrancier de ce qui a déjà été réalisé dans l'exagération* ».
3. Helen Keller (1880-1967) était sourde, muette et aveugle de naissance mais a réussi à obtenir des diplômes universitaires et à devenir écrivaine.
4. Sur cette interprétation du test de Turing, voir Lestel 1996 et D. Lestel, L. Bec et J. L. Lemoigne (1993).
5. Dans une postface à son livre *I and Thou*, Buber fait une timide ouverture en direction de l'animal, mais il ne va pas très loin.
6. Je préfère la notion de « vraisemblance » à celle de « réalisme » parce que la première est une notion clairement relationnelle alors que la deuxième est une notion ontologique qui implique déjà d'avoir répondu à la question de savoir ce qui est réel, question qui fait justement partie du problème.
7. Il existe tout un courant de psychologie cognitive qui considère que la robotique est très importante pour étudier comment les enfants fonctionnent. De mon point de vue, c'est une perte de temps qui repose sur des erreurs de raisonnement.
8. C'est un point qu'ont très bien compris Emmanuel Grimaud et Zaven Paré (2011).
9. Le psychanalyste Jean-Jacques Lecercle montre que la création du monstre par Frankenstein est l'analogue d'un acte sexuel (Lecercle, 1988). Et l'un des films sur le monstre de Frankenstein qui a eu le plus de succès est précisément le second film de Whale, *The Bride of Frankenstein*, en 1935.
10. "Existence" est pris dans ce texte au sens existentialiste du terme plus que dans son sens ontologique.
11. Sur la notion de communauté hybride de partage de sens, d'intérêts et d'affects, voir Lestel 1996.
12. Voir à ce sujet le livre intrigant d'Erik Davis, 1998, *Technognosis : myth, magic + mysticism in the age of information*, New York, Three Rivers Press. Thierry Bardini a évoqué cet aspect des technologies modernes de l'information dans Bardini 2000 et sa reprise dans Fred Turner 2006.
13. Les taoïstes chinois avaient un terme pour une telle passivité positive : *wu wei*.
14. Le Paro est un robot de compagnie très populaire au Japon qui a la forme d'un petit phoque.

15. L'Aibo, construit par Sony avec une forme de petit chien, a été le plus populaire des robots de compagnie.

---

## RÉSUMÉS

À partir de la deuxième moitié du XX<sup>e</sup> siècle, un ensemble croissant de créatures se développent moins selon des processus évolutionnistes classiques qu'en fonction de compromissions perturbantes et d'irritations fécondes avec les technologies humaines. En émerge entre autres choses une vision des êtres vivants à la deuxième personne très différente des biologies à la troisième personne que nous connaissions jusque-là. D'où la possibilité d'adopter une pluralité de théories du vivant, plus ou moins compatibles les unes avec les autres, mais susceptibles de cohabiter sans s'agresser mutuellement pour autant. Une façon d'assumer le fait que le vivant est un phénomène qui déborde toujours les théories qu'on peut en faire.

## AUTEUR

### DOMINIQUE LESTEL

Dominique Lestel est un philosophe qui travaille depuis vingt ans sur les intoxications conjointes de l'humain et du non-humain. Il a publié de nombreux livres sur le sujet, en particulier *L'animal est l'avenir de l'homme* (Fayard, 2010), *Apologie du carnivore* (Fayard, 2011) et *The friends of my friends : a new look at human/animal relationships* (Columbia University Press, 2015). La revue anglaise *Angelaki: Journal of the Theoretical Humanities* a consacré un numéro spécial à son travail en 2014.

[lestel [at] ens.fr]